UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/703,174	10/31/2000	Charu C. Aggarwal	YOR920000430US1	7445

	7590	12/23/2005

William E Lewis
Ryan Mason & Lewis LLP
90 Forest Avenue
Locust Valley, NY 11560

EXAMINER
HILLERY, NATHAN

ART UNIT	PAPER NUMBER
2176	

DATE MAILED: 12/23/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *28 September 2005*.

2a)☒ This action is **FINAL**.   2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-27* is/are pending in the application.

　　4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-27* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

　　Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

　　Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

　　a)☐ All   b)☐ Some * c)☐ None of:

　　　1.☐ Certified copies of the priority documents have been received.

　　　2.☐ Certified copies of the priority documents have been received in Application No. _____.

　　　3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

　　* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
　　Paper No(s)/Mail Date _____.

4) ☐ Interview Summary (PTO-413)
　　Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: _____.

DETAILED ACTION

1.   This action is responsive to communications: Amendment filed on 9/28/05.

2.   Claims 1 – 27 are pending in the case. Claims 1, 10, and 19 are independent.

3.   The rejection of claims 1 – 27 under 35 U.S.C. 102(e) as being anticipatory has

been maintained.

Claim Rejections - 35 USC § 112

4.   The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

5.   Claims 1 – 27 are rejected under 35 U.S.C. 112, first paragraph, because the

specification, while being enabling for *not assuming* *a specific model for web linkage*

structure (Specification, p 4, line 24), does not reasonably provide enablement for *not*

assuming a predefined link structure. The specification does not enable any person

skilled in the art to which it pertains, or with which it is most nearly connected, to make

and use the invention commensurate in scope with these claims. Not only is the

specification not enabling for not assuming a predefined link structure but also the

claims seem to be inoperative. Specifically, by definition an information network is a

predefined link structure as is evidenced by the cited NPL reference of Hofmann, which

teaches that *In the World Wide Web, myriads of hyperlinks connect documents and*

pages to create an unprecedented, highly complex graph structure - the Web graph (p

369, left column, lines 1 – 3). Consequently, the claim limitation, *without assuming a*

predefined link structure, will not be considered further on the merits at this time.

6.   The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly
> claiming the subject matter which the applicant regards as his invention.

7.   Claims 1 – 27 are rejected under 35 U.S.C. 112, second paragraph, as being

indefinite for failing to particularly point out and distinctly claim the subject matter which

applicant regards as the invention.

8.   **Regarding independent claims 1, 10, and 19**, it is unclear what Applicant

means by "such that one or more documents are retrieved from the information network

that satisfy the user-defined predicate without assuming a predefined link structure",

since it appears to be inoperative and impossible; consequently, the metes and bounds

of the claim are unclear.

9.   **Regarding dependent claims 2 – 9, 11 – 18 and 20 – 27**, the claims are

rejected for fully incorporate the deficiencies of the base claim(s) from which they

depend.

Claim Rejections - 35 USC § 102

1.   The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that

form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by
> another filed in the United States before the invention by the applicant for patent or (2) a patent
> granted on an application for patent by another filed in the United States before the invention by the
> applicant for patent, except that an international application filed under the treaty defined in section
> 351(a) shall have the effects for purposes of this subsection of an application filed in the United States
> only if the international application designated the United States and was published under Article 21(2)
> of such treaty in the English language.

2.   Claims 1 – 27 are rejected under 35 U.S.C. 102(e) as being anticipated by

Bharat et al. (US006112203A).

3.   **Regarding independent claim 1**, Bharat et al. teach that *in one aspect of the*

invention, the documents are Web pages connected to each other by hyperlinks. The

identities of the documents, and the hyperlinks are in the form of a string called a

Uniform Resource Locator (URL). The URLs specify the addresses of the various

documents. The set of documents can be produced by combining the set of results

from a Web search engine in response to a user query (which we call the `start-set`),

with pages that either link to or are linked from the start-set documents. Terms of the

query imply a topic of interest on which the user requested the search to be made. The

nodes in the start set are first scored according to their connectivity, and the number of

terms of the query that appear as unique sub-strings in the URL of the represented

documents. The score is a weighted sum of the number of directed edges to and from

a node and the number of unique sub-strings of the URL that match a query term

(Column 2, line 66 – Column 3, line 15), which provide for **retrieving one or more**

documents from the information network that satisfy a user-defined predicate.

Bharat et al. teach that *in our present invention, we use only a subset of the pages for*

the purpose of content analysis. The subset of influential pages are selected by a

heuristic that is based on the URLs of the pages in the start set 201 and their

connectivity. This information can be determine from the graph 211 without having to

fetch the pages themselves. The heuristic selects nodes based on "in-degree," i.e., the

number of edges 213 pointing at a node, "out-degree" (out-going edges) and

comparison of the key words in the query with unique sub-strings of the URL.

Specifically, in step 220, we score each page p of the input set 201 to determine a value

Score(p) 225. Let n_p be the node representing page p. The score is determined by:

*Score(p) = in_degree + 2 * (num_query_matches) + out_degree, where in_degree is the*

number of edges pointing at node n_p, num_query_matches is the number of unique sub-

strings of the URL of the page p that exactly match a term in the user's query, and out-

degree is 1 if the node n_p has at least one edge pointing to another page; otherwise, the

value of out-degree is 0. Note, the values Score(p) 225 can be determined without

having to fetch the actual pages (Column 5, lines 47 – 67), which provide for **collecting**

statistical information about the one or more retrieved documents as the one or

more retrieved documents are analyzed. Bharat et al. teach that *in step 230, a small*

subset of start set pages 235 with the highest values n_p are selected. We select thirty,

although it should be understood, that other sized subsets can also be used. The subset

of pages 235 is used to distill the broader query topic Q 245 in step 240. Each page of

the subset 235 is fetched, and the first, for example, one-thousand words of all of the

selected pages are concatenated to form Q (Column 6, lines 1 – 8), which provide for

using the collected statistical information to automatically determine further

document retrieval operations such that one or more documents are retrieved

from the information network that satisfy the user-defined predicate.

4.   **Regarding dependent claim 2**, Bharat et al. teach that *in response to a query*

composed by a user, the search engine returns a result set which satisfies the terms

(key words) of the query (Column 4, lines 11 – 14), which provides that **the user-**

defined predicate specifies content associated with a document.

5.   **Regarding dependent claims 3 and 4,** Bharat et al. teach that *during a content*

analysis phase, a relevance weight is assigned to a carefully chosen subset of the

nodes in the graph. The relevance weights are based on the similarity of each

represented document to the distilled topic as determined above. The relevance weight

of a document is further increased when the document includes words that are terms of

the query (Column 3, lines 21 – 27), which provide that **the statistical information**

collection step uses content of the one or more retrieved documents and that **the**

statistical information collection step considers whether the user-defined

predicate has been satisfied by the one or more retrieved documents.

6.   **Regarding dependent claims 5 and 6,** Bharat et al. teach that *in step 260, we*

assign a similarity weight to each node 213 of the sub-graph 255. Various document

similarity measuring techniques have been developed in Information Retrieval to

determine the goodness of fit between a "target" document and a collection of

documents. These techniques typically measure a similarity score based on word

frequencies in the collection and a target document (Column 6, lines 51 – 57), which

provide that **the collected statistical information is used to direct further document**

retrieval operations toward documents which are similar to the one or more

retrieved documents that also satisfy the predicate, and that **the collected**

statistical information is used to direct further document retrieval operations

toward documents which are more likely to satisfy the predicate than would

otherwise occur with respect to document retrieval operations that are not

directed using the collected statistical information.

7.   **Regarding dependent claim 7**, Bharat et al. teach that *in one prior art*

technique, an algorithm for connectivity analysis of a neighborhood graph (n-graph) is

described by Kleinberg ...The algorithm analyzes the link structure, or connectivity of

Web pages "in the vicinity" of the result set to suggest useful pages in the context of the

search that was performed (Column 1, lines 55 – 64), which provide the capability that

the collected statistical information is used to direct further document retrieval

operations toward documents which are linked to by other documents which also

satisfy the predicate.

8.   **Regarding dependent claim 8**, Bharat et al. teach that *FIG. 1 shows a*

distributed network of computers 100 that can use our invention. Client computers 110

and server computers 120 (hosts) are connected to each other by a network 130, for

example, the Internet. The network 130 includes an application level interface called the

World Wide Web (the "Web") (Column 3, lines 59 – 64) and that *although the invention*

is described with respect to documents that are Web pages, it should be understood

that the invention can also be worked with any linked data objects of a database whose

content and connectivity can be characterized (Column 4, lines 4 – 8), which provide for

the information network is the World Wide Web and a document is a web page.

9.   **Regarding dependent claim 9**, Bharat et al. teach that *in our present invention,*

we use only a subset of the pages for the purpose of content analysis. The subset of

influential pages is selected by a heuristic that is based on the URLs of the pages in the

start set 201 and their connectivity. This information can be determined from the graph

211 without having to fetch the pages themselves. The heuristic selects nodes based

on "in-degree," i.e., the number of edges 213 pointing at a node, "out-degree" (out-going edges) and comparison of the key words in the query with unique sub-strings of the URL (Column 5, lines 47 - 56), which provides that **the statistical information collection step uses one or more uniform resource locator tokens in the one or more retrieved web pages.**

10.   **Regarding claims 10 – 27**, the claims incorporate substantially similar subject matter as claims 1 – 9, and are rejected along the same rationale.

Response to Arguments

11.   Applicant's arguments filed 9/28/05 have been fully considered but they are not persuasive.

12.   In response to Applicant's arguments that Bharat fails to teach, disclose, or suggest **using the collected statistical information to automatically determine further document retrieval operations such that one or more documents are retrieved from the information network that satisfy the user-defined predicate without assuming a predefined link structure,** it should be noted that this particular limitation was addressed in the rejection of the claims under 35 USC 112, first and second paragraphs above. It should also be noted that the Office can see no difference in Bharat and the claimed invention as explained by Applicant. Specifically, by applicant's own admission, Bharat requires a preconstructed graph that includes nodes and directed edges, where each node represents a document and the directed edges represent the links connecting the documents. This is what Bharat refers to as a "start-set". Thus, Bharat assumes a predefined link structure. However, while the method of

the invention may generate a graph-like structure as it "crawls," it merely starts off with a start list, which is merely a list of Uniform Resource Locators (URLs) (p 8, last two paragraphs). The Office interprets Bharat's start-set as being synonymous with the present invention's start list. Furthermore, as described in the rejection of the claims under 35 USC 102(e), the Office maintains that Bharat anticipates the claimed invention.

13.   In response to Applicant's arguments that Bharat fials to teach, disclose, or suggest **collecting statistical information about the one or more retrieved documents as the one or more retrieved documents are analyzed**, it should be noted that the Office interprets "statistical information" to be any information that relates to a statistic. Since the Merriam Webster Dictionary defines a statistic as a quantity (as the mean of a sample) that is computed from a sample, the Office has interpreted the $Score(p)$ computed by Bharat (Column 5, lines 47 – 67) as explained above in the rejection of claims under 35 USC 102(e) to be synonymous with statistical information and therefore teaches the claim limitation.

Conclusion

14.   Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL.** See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not

mailed until after the end of the THREE-MONTH shortened statutory period, then the

shortened statutory period will expire on the date the advisory action is mailed, and any

extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of

the advisory action. In no event, however, will the statutory period for reply expire later

than SIX MONTHS from the date of this final action.

Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Nathan Hillery whose telephone number is (571) 272-

4091. The examiner can normally be reached on M - F, 10:30 a.m. - 7:00 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Heather R. Herndon can be reached on (571) 272-4136. The fax phone

number for the organization where this application or proceeding is assigned is 571-

273-8300.

Information regarding the status of an application may be obtained from the

Patent Application Information Retrieval (PAIR) system. Status information for

published applications may be obtained from either Private PAIR or Public PAIR.

Status information for unpublished applications is available through Private PAIR only.

For more information about the PAIR system, see http://pair-direct.uspto.gov. Should

you have questions on access to the Private PAIR system, contact the Electronic

Business Center (EBC) at 866-217-9197 (toll-free).

WILLIAM BASHORE
PRIMARY EXAMINER
12/21/2005

NH